



Vocabulary learning and retention through multimedia glossing

Nasrin Ramezanali, *University of Western Ontario*

Farahnaz Faez, *University of Western Ontario*

Abstract

Drawing on Mayer's (2014) cognitive theory of multimedia learning, the purpose of this study was to examine which modes of gloss presentation (i.e., L2 definition, aural, and video animation) are effective for learners' vocabulary learning and delayed word recollection. One control group and three experimental groups were formed by 132 intermediate language learners. Pre- and post-tests of productive recall and multiple-choice productive recognition were administered, and learners' perceptions toward glossing were examined through a questionnaire and interviews. Quantitative data were analyzed using ANCOVA, and themes that emerged from the qualitative data were identified. The quantitative findings revealed that dual glossing modes were more effective than single glossing modes for many test sessions. However, single glossing was also effective for a few test sessions. The questionnaire and interview data showed that learners preferred the dual glossing mode of L2 definition and video animation. The findings provide insights for vocabulary learning and teaching.

Keywords: *Glossing, Vocabulary Learning and Teaching, Immediate and Delayed Word Retention, Multimedia Learning*

Language(s) Learned in This Study: *English*

APA Citation: Ramezanali, N., & Faez, F. (2019). Vocabulary learning and retention through multimedia glossing. *Language Learning & Technology*, 23(2), 105–124. <https://doi.org/10.125/44685>

Introduction

Vocabulary development is one of the most important aspects of foreign language (FL) or second language (L2) learning and teaching. Common forms of vocabulary teaching include word-lists, dictionary use, vocabulary cards, negotiating vocabulary meanings, and glosses. Glosses are short definitions in a learner's first language (L1) or L2 that accompany text to clarify the meaning of unknown words (Nation, 2013). Glosses facilitate comprehension of a text (Khezrlou & Ellis, 2017), save a learner's time and effort (Nation, 2001), help avoid wrong inferences (Nation, 2013), and are therefore potentially effective for vocabulary learning (Webb, 2010). While several studies have reported that glosses facilitate vocabulary learning (e.g., Jacobs, Du Fon, & Hong, 1994; Jung, 2016; Rassaei, 2017), there are inconsistencies reported on the efficacy of different gloss modes and the superiority of single versus dual gloss modes. In a series of articles, Chun and Plass (1996a, 1996b) reported that the glossing of text and picture was more effective than text and video. However, Al-Seghayer (2001) and Lin and Tseng (2012) found the opposite, and still other studies reported an equal but positive effectiveness of the two modes (Akbulut, 2007; Mohsen & Balakumar, 2011). Although several studies have attested to the superiority of dual glosses over single text-only glosses (e.g., Al-Seghayer, 2001; Chun & Plass, 1996a, 1996b), others (e.g., Acha, 2009; Boers, Warren, He, & Deconinck, 2017) have reported no benefit to additional gloss modes. In fact, Ariew and Erçetin (2004) identified a distracting effect for video inclusion, and other studies also found a distracting effect on the addition of utterance to the glossed words (Kim & Gilman, 2008) and no significant difference for L2 vocabulary learning (Yeh & Wang, 2003). Hence, this study aimed to examine the effectiveness of simultaneous multimedia learning tools such as textual, aural, and video animation glossing on L2 learners'

word learning and delayed word retention.

Theoretical Framework

This research draws on Mayer's (2014) cognitive theory of multimedia learning (CMTL) that accounts for the value and effect of multimedia presentations in language learning environments. CTML is an instructional theory focusing on how people learn from multimedia presentations (Mayer, 2014). CTML emphasizes that learners attempt to build meaningful connections between words and pictures and process them actively in long-term memory. The theory is based on three assumptions: dual channels, limited capacity, and active processing. The dual channel assumption suggests that humans have two separate information processing channels: auditory or verbal and visual or pictorial. The verbal channel receives information (e.g., spoken or written words, narrations, and sounds) through the ear and eyes, and the visual channel processes information (e.g., pictures, graphs, video animation clips, and on-screen texts) through the eyes. The limited capacity assumption signifies that there is a limited amount of information that learners can process in each channel at one time. Therefore, the information needs to be organized into two separate mental models. Words are stored in a verbal mental model and images in a visual mental model. Finally, to build a connection between verbal and visual representations and integrate them into the learner's existing knowledge, some cognitive activities should be processed in the long-term memory, accounting for the third assumption (i.e., active processing).

The theory implies that L2 learners' interactivity with multimedia input is best facilitated and enhanced by connecting written and pictorial cues to both visual and verbal systems rather than to just a single mode. Furthermore, utilizing multimedia annotations to teach unfamiliar target words can help learners to experience word learning through the cognitive processes of selecting relevant words and images, organizing them into verbal and visual representations, and finally integrating the words into corresponding verbal and visual stimulus (Mayer, 2014). To minimize the chances of overloading a learners' cognitive system, both verbal and visual materials can be displayed simultaneously rather than successively (Mayer & Fiorella, 2014).

Vocabulary Learning with Multimedia Glossing

Mayer (2014) defines multimedia as the presentation of both words and pictures to language learners. In his definition, words signify any material presented in verbal form such as spoken texts or printed texts, and pictures represent static or dynamic pictorial materials such as illustrations, graphs, diagrams, maps, animations, or videos. Thus, multimedia learning refers to "building mental representations from words and pictures" (p. 2).

Multimedia glossing emerged because of advancements in computer technology and media, and the integration of glosses into software programs to improve L2 vocabulary learning. In hypermedia environments, multimedia glosses take the form of various vocabulary annotations such as texts, audio, pictures, and video animation (AbuSeileek, 2011; Akbulut, 2007; Al-Seghayer, 2001; Chun & Plass, 1996a, 1996b; Khezrlou & Ellis, 2017; Rusanganwa, 2015; Yanguas, 2009; Yeh & Wang, 2003). Multimedia glossing presents the target words via multiple types of resources in a multimedia-based learning setting such as a language lab or a classroom equipped with a computer, screen, projector, internet, or other technological tool. Given the positive role of multimedia instruction in language classes, there is a tendency to view computer-based glossing as an interactive, context-based, and learner-oriented strategy to facilitate L2 vocabulary learning.

There are several instructional and educational advantages to the application of multimedia glossing, such as (a) providing L2 learners with adequate use of authentic texts, (b) attending to learners' preferred modes of vocabulary learning and learning styles, (c) improving vocabulary retention, and (d) saving learners' time and effort in reading L2 texts. Furthermore, in favor of computer-based multimedia glossing for vocabulary learning and retention, the results of three meta-analyses (Abraham, 2008; Vahedi, Ghonsooly,

& Pishghadam, 2016; Yun, 2011) found that multimedia glossing had a large positive effect size (Abraham, 2008).

Glosses can fall into single, dual, and multimodal types. The single gloss types include text-only, audio-only, picture-only, and video-animation-only glosses (Boers, Warren, He, & Deconinck, 2017; Rassaei, 2017). Dual gloss mode presents information through the combination of text and pictures (Chun & Plass, 1996a, 1996b), text and audio (Yeh & Wang, 2003), or text and videos (Al-Seghayer, 2001). A multimodal gloss type can also refer to short definitions or explanations in computerized or non-computerized texts in a range of different glossing conditions. Some studies have examined the efficacy of glossing modes in paper-and-pencil format (Boers, Warren, He, & Deconinck, 2017) and others have investigated this impact in a multimedia setting (Al-Seghayer, 2001).

Considerable research has been conducted on the use of different combinations of multimedia glossing (i.e., single vs. dual). Chun and Plass (1996a, 1996b) conducted a series of studies that investigated the effectiveness of dual glossing modes of verbal and visual modes on the vocabulary learning of L2 learners enrolled in a German course. The results revealed that participants preferred visual modes and acquired more words with annotations of text and picture than text only or text and video. The study supported the efficacy of dual presentations over a single condition.

Guided by the question of which gloss combination, video clips or still images, was more effective for facilitating vocabulary acquisition, Al-Seghayer (2001) conducted a within-participant design study among learners of English as a second language (ESL) in a multimedia learning environment. The findings of both recognition and production vocabulary tests revealed that words glossed with text and picture and those with text and video were learned better than those with text-only glosses. However, in contrast to the studies by Chun and Plass (1996a, 1996b), the participants recalled more words when video clips were provided than when still pictures were shown. Long-term retention still needs further investigation.

Lin and Tseng (2012) conducted a quasi-experimental research study with beginning L2 learners using glossing modes in the learners' L1 to define a target vocabulary (i.e., L1 definition alone, L1 definition and pictures, and L1 definition and video). The findings revealed that learning difficult words with textual definitions and videos was more effective than learning with textual definitions and pictures or textual definitions alone. To further confirm the efficacy of dual over single glossing modes, Yanguas (2009) investigated the effectiveness of multimedia glosses on text comprehension and vocabulary learning via a computerized text using four gloss conditions: textual, pictorial, textual and pictorial, and no gloss. The findings showed that the group with the combination or dual modes outperformed the other groups.

Employing a rather different gloss combination, Yeh and Wang (2003) also investigated the effectiveness of text annotation only, text and picture annotations, and text, picture, and sound annotations in association with FL learners' vocabulary learning in a multimedia setting. Of note, this study was different from others (e.g., Yoshii & Flaitz, 2002; Kost, Foss, & Lenzini, 1999) as it added sound as a component to multimedia glosses and both the L1 and the L2 were incorporated in the textual glosses. The findings, falling in line with the premises of CTML, indicated that the combination of text and picture was the most effective type of vocabulary glossing. The study revealed, however, that the addition of aural information lowered vocabulary gains. The researchers attributed the lower vocabulary gains to the potential overloading of verbal and visual information, exceeding the participants' cognitive processing capacities; inadequate time for the participants to complete the task; and the possibility of different learning styles.

Though the above-mentioned studies support the positive effects of dual glossing on L2 vocabulary acquisition, there are conflicting findings in the field. For example, Akbulut's (2007) findings indicated no significant differences between the two visual modes of text and pictures and text and videos. Boers, Warren, He, and Deconinck (2017) also reported that textual and pictorial gloss modes did not help learners learn and retain words any better than text-only explanations. In a discussion article, Boers, Warren, Grimshaw, and Siyanova-Chanturia (2017) argued that the inclusion of two annotation types (verbal and visual) invited "two look-ups" (p. 6), while one gloss annotation (textual definition alone or visual

representation alone) required learners to look at the annotated word only once. The amount of attention to an unfamiliar word is a significant predictor for word learning and retention and the application of multimodality is just one of many possible ways of promoting longer attention. Therefore, the learning and retention of words was not necessarily dependent on the multimodality of annotations. This explanation might lend support to studies that report the outperformance of single-mode annotations over dual modes for word learning and retention. Overall, a survey of the literature shows inconsistencies regarding multimedia glossing and which gloss types and combinations are more effective in facilitating L2 vocabulary learning.

Research Purpose and Research Questions

To address the identified gap, the present study examined the effectiveness of multimedia glossing modes (L2 definition alone, L2 definition and audio glossing, and L2 definition and video animation glossing) on L2 learners' vocabulary learning and retention. The following research questions guided the study:

1. Which glossing modes contribute significantly to L2 learners' immediate and delayed word learning and retention?
2. What are L2 learners' attitudes toward different glossing modes? Which mode or modes do they prefer, and why?

Methodology

This study employed a quasi-experimental mixed-methods research with a between-participant design to compare the performance of four groups of participants (one control group and three experimental groups) that received different glossing modes. A questionnaire and semi-structured interviews were utilized to examine learners' preferences and perceptions toward glossing modes.

Research Instruments

Several instruments were utilized to gather the data. First, a demographic information form was used to understand participants' backgrounds.

Next, a 2000-word level vocabulary levels test (VLT) was used to control for homogeneity and proficiency differences (Schmitt, Schmitt, & Clapham, 2001).

Three different reading passages were employed in three different glossing modes (hyperlinked with L2 definition alone, L2 definition and audio glossing, and L2 definition and video animation glossing; see [Appendix A](#)). The texts were checked against the Flesch-Kincaid grade level readability formula and content appropriateness. Each text was 300–400 words in length, and the lexical coverage was between 95% and 98% (Nation, 2013; for samples of the reading excerpts, see [Appendix B](#), [Appendix C](#), and [Appendix D](#)).

A total of 33 target words were glossed (for word samples, see [Appendix E](#)). The following criteria were considered for word inclusion: (a) A 70% criterion was set for target word inclusion, meaning if 70% of the participants did not know the word, the word was considered a target glossed word (Perez, Peters, Clarebout, and Desmet, 2014). (b) Only concrete nouns which had simple and clear textual definitions and appropriate video animation clips were selected. (c) All target words and their definitions were controlled for consistency with the British National Corpus and the Corpus of Contemporary American English (BNC–COCA-25) lexical vocabulary profile to ensure that they were within the 2000-word level. (d) Only 2% to 5% of words in a 300- to 400-word passage were glossed, following Nation's (2001) guide of gloss coverage. Applying the above criteria, 12 words were glossed in Passage 1, 11 words in Passage 2, and 10 words in Passage 3. All target words were bold-faced, underlined, and hyperlinked.

Two vocabulary measurement tests of productive recall (PR) and multiple-choice (MC) productive recognition tests were administered. The tests served as pre-tests, immediate post-tests, and delayed post-

tests. In the PR test, learners were asked to recall each target word based on its definition in the stem. In the MC test, each target word was written, and the participants were required to choose the correct alternative among four items (the correct item and three distractors) for each definition in the stem (for sample questions, see [Appendix F](#) and [Appendix G](#)). The pre-tests measured the number of words that were unknown to the participants. The immediate and delayed post-tests administered after instruction were used to gauge learners' immediate word knowledge and delayed word retention, respectively. The test items were piloted before the actual implementation and were checked for internal consistency using Cronbach alpha (α). The binominal scoring of 1 (correct) and 0 (incorrect) was applied for MC tests. The trinomial scoring of 2 (correct), 1 (partially correct), and 0 (incorrect) was employed for the PR test. The responses that contained minor spelling mistakes such as “bedbog” for *bedbug* or “rudent” for *rodent*, were regarded as a sign of learning on behalf of the participants and were thus counted as partially correct responses while scoring.

A questionnaire consisting of nine closed-ended questions focused on which glossing modes helped participants (a) to learn the words easily, (b) to retain them longer, and (c) to ascertain whether participants were willing to use the modes for future practice. Each question was repeated three times for each glossing mode, and participants were asked to rank their responses on a 5-point Likert scale (1 = strongly agree; 2 = agree; 3 = neither agree nor disagree; 4 = disagree; 5 = strongly disagree).

A final interview was conducted to gain a deeper understanding of which modes of glossing the participants preferred and why. The interview questions were self-developed and piloted (see [Appendix H](#)).

Research Participants

Participants included 132 adult intermediate learners of English—male ($n = 50$) and female ($n = 82$)—ranging from 16 to 25 years of age with an average age of 17.56. All participants, except one, spoke Farsi (Iran's official language) as their L1 and were learning English as their L2. Only one participant knew additional languages (Farsi, Turkish, and Kurdish). The participants attended the institute's language classes three times a week and each instructional session lasted 90 minutes. The research participants were situated in four intact intermediate classes (24–39 students per class). Accordingly, there were four groups in the study: a control group (Group A) and three experimental groups (Groups B, C, and D).

The study was conducted at a private language institute in Iran, and one of the institute's administrative staff helped to arrange the classes for the study. Though the researchers were not regular instructors of the institute, the first author instructed the intervention related to this study. Data was collected in Fall 2016.

Data Analysis

The researchers conducted an analysis of covariance (ANCOVA) using SPSS 20.0 software (4 [four groups] \times 2 [two immediate post-tests] \times 2 [two delayed post-tests]). The level of significance was set at an alpha level of $p < .05$. For the analyses where statistical test assumptions were not met, the alpha level was set at $p < .025$ (Tabachnick & Fidelle, 2013). The independent variables included different glossing modes of instruction: text-effect (using the same passage for each group on each instructional day) and test effect (administering the PR before the MC test). The dependent variable included all test scores on immediate and delayed post-tests. The findings of the questionnaire were analyzed and converted into percentages. The interviews were conducted by the first author, audio-recorded, and then fully transcribed. The transcripts were checked twice to ascertain that the statements were accurate. Manual coding was used to code and extract the themes that recursively emerged throughout the document. The most frequent themes were underlined and kept for analysis (inductive coding). Only themes pertaining to the focus of this paper are discussed.

Research Procedures

Phase 1. Selecting Target Glossed Words

Based on two pilot studies that identified typically unfamiliar words for intermediate students in Canada

and Iran, the researchers selected 33 words. In the piloting stage, 12–14 intermediate adult ESL or English as a foreign language (EFL) volunteer participants whose profiles were similar to the target participants were asked to read three English passages and underline the words that they did not know. The criteria previously outlined in this article were used to select the target words.

Phase 2. Pre- and Post-Test Research Procedures

The research was conducted over five weeks. In Week 1, during the pre-instructional session, the participants signed a consent form to participate and were given the VLTs and two pre-tests (PR and MC tests) to complete. They were then assigned to one of four groups (a control group or one of three experimental groups) from their intact classes. In Week 2, during instructional sessions, the participants of the experimental groups were presented with the target glossed words via the three glossing modes. The instruction was counter-balanced for the experimental groups and each text was glossed in all three modes and used every other day for three days (see [Table 1](#)).

Table 1. *Research Procedure for Week 2 Instruction*

Group	Session 1	Session 2	Session 3
A	Text 1 (Control)	Text 2 (Control)	Text 3 (Control)
B	Text 1 (L2 definition alone)	Text 2 (L2 definition and audio)	Text 3 (L2 definition and video animation)
C	Text 1 (L2 definition and audio)	Text 2 (L2 definition and video animation)	Text 3 (L2 definition alone)
D	Text 1 (L2 definition and video animation)	Text 2 (L2 definition alone)	Text 3 (L2 definition and audio)

Group A received no specific intervention or external enhancement. The instructor read aloud each passage from a large screen monitor, pausing for 5–7 seconds after each glossed word. This reading presentation resembled a common form of instruction in Iran where the teacher reads aloud a passage and the students follow along on a large screen. This practice is somewhat different from some other studies (e.g., Al-Seghayer, 2001; Yanguas, 2009) where participants were asked to read a passage in a multimedia program on individual computers. Because of technological limitations, providing access to individual computers for each student was not feasible.

After each day's instruction, the participants of all groups were given two immediate PR and MC post-tests. There was no activity during Week 3. Two weeks later (Week 4), the participants were asked to complete the PR and MC vocabulary post-tests. The items of the delayed post-tests were reordered to avoid guessing. Given the time constraints of this study regarding access to participants, a more long-term post-test (more than two weeks) was not feasible. In all sessions, the PR test was given before the MC test to prevent test effect and learning effect. From 132 participants, only 83 learners attended this final session. The tests were administered in pencil-and-paper format.

Phase 3. Questionnaire and Interview Research Procedures

Phase 3 of the research procedure conducted in two parts. After the delayed post-tests, the first author administered a pencil-and-paper questionnaire to the participants of the experimental groups to examine their perceptions toward the three glossing modes. Then, the first author interviewed nine participants from the experimental groups who volunteered to further elaborate on which modes of glossing they preferred and why. The 30-minute interviews were conducted during Week 5.

Results

The following analyses were performed: (a) a homogeneity check of the participants using the VLT, (b) the

between-participant comparisons to examine which glossing modes contributed significantly to vocabulary learning and retention in immediate and delayed vocabulary tests (Research Question 1), and (c) the questionnaire and interview data (Research Question 2).

Homogeneity Check

Table 2 presents the descriptive statistics on the performance of all participants on the VLT. As shown, the mean scores of the four groups were relatively similar.

Table 2. *Descriptive Results of the VLT*

Group	<i>N</i>	<i>M</i>	<i>SD</i>
Group A (Control)	24	23.45	3.93
Group B	39	23.38	3.39
Group C	36	24.41	1.91
Group D	33	24.51	1.90

To ensure the homogeneity of the participants with respect to the VLT, the Kruskal Wallis Test was used. The findings showed that groups were not significantly different from one another in terms of the VLT mean scores ($\chi^2 = 2.71$, $df = 3$, $p = .430$) and were thus homogenous.

Impact of Different Glossing Modes

To address Research Question 1, participants' pre-test scores were compared to immediate post-test scores for immediate effect, and then pre-test scores were compared to delayed post-test scores for delayed effect, using ANCOVA. For conciseness in the analyses, sub-tests of PR and MC were created and all ANCOVA assumptions were ensured.

Different Glossing Modes and Immediate Word Retention

With respect to the performance of participants on the PR immediate post-test, the findings of ANCOVA revealed that all experimental groups were statistically different from the control group. Therefore, glossing was effective for immediate word learning. To find out which glossing mode was more effective, a pairwise comparison was conducted (see Table 3). It showed that the three experimental groups achieved significantly higher means than the control group ($p < .05$) on all immediate post-tests.

In the Immediate Sub-Test 1 comparison, Groups C and D performed significantly better than Group B ($p < .05$), whereas Group B (single-mode) was only slightly better than Group A ($p < .05$). Thus, dual glossing modes were significantly more effective than the single or no glossing modes for learners' immediate word learning. In the Immediate Sub-Test 2 comparison, Group C which received words through L2 definition and video animation performed significantly better than Group B ($p < .05$). Finally, in the Immediate Sub-Test 3 comparison, there was no significant difference among the performances of all the experimental groups ($p > .05$).

Table 3. Retention for Immediate PR Tests

Group	Comparison	Immediate Sub-Test 1		Immediate Sub-Test 2		Immediate Sub-Test 3	
		Difference (M)	p	Difference (M)	p	Difference (M)	p
Group A	Group B	-4.57*	.009	-3.71*	.025	-4.42*	.004
	Group C	-8.05*	.000	-7.10*	.000	-4.89*	.001
	Group D	-10.57*	.000	-4.86*	.001	-4.72*	.002
Group B	Group A	4.57*	.009	3.71*	.025	4.42*	.004
	Group C	-3.48*	.047	-3.38*	.034	-0.47	.999
	Group D	-6.00*	.000	-1.15	.919	-0.30	1.000
Group C	Group A	8.05*	.000	7.10*	.000	4.89*	.001
	Group B	3.48*	.047	3.38*	.034	0.47	.999
	Group D	-2.52	.261	2.23	.303	0.17	1.000
Group D	Group A	10.57*	.000	4.86*	.001	4.72*	.002
	Group B	6.00*	.000	1.15	.919	0.30	1.000
	Group C	2.52	.261	-2.23	.303	-0.17	1.000

*The mean difference was significant at $p < .05$

In the MC test, ANCOVA results showed that there were statistically significant differences among the experimental groups in comparison to the control group with large effect sizes, ranging from .15 to .36. However, unlike the PR test results, the three experimental groups were not significantly different from one another ($p > .05$) in immediate word retention (see pairwise comparison in Table 4). As an effect size measure, partial eta squared (ηp^2) was used.

Table 4. Retention for Immediate MC Tests

Group	Comparison	Immediate Sub-Test 1		Immediate Sub-Test 2		Immediate Sub-Test 3	
		Difference (M)	p	Difference (M)	p	Difference (M)	p
Group A	Group B	-4.53*	.000	-2.93*	.000	-2.08*	.002
	Group C	-4.01*	.000	-2.57*	.000	-1.65*	.023
	Group D	-4.94*	.000	-3.63*	.000	-2.29*	.000
Group B	Group A	4.53*	.000	2.93*	.000	2.08*	.002
	Group C	0.52	.952	0.35	.986	0.43	.956
	Group D	-0.40	.987	-0.70	.719	-0.21	.999
Group C	Group A	4.01*	.000	2.57*	.000	1.65*	.023
	Group B	-0.52	.952	-0.35	.986	-0.43	.956
	Group D	-0.93	.575	-1.05	.275	-0.64	.766
Group D	Group A	4.94*	.000	3.63*	.000	2.29*	.000
	Group B	0.40	.987	0.70	.719	0.21	.999
	Group C	0.93	.575	1.05	.275	0.64	.766

*The mean difference was significant at $p < .05$

Table 5 summarizes the findings of glossing modes in immediate retention for both the PR and the MC tests. In summary, based on the PR test results, dual glossing modes were effective for two out of the three test sessions for immediate word learning (Days 1 and 2), All three glossing modes were equally effective for one out of three test sessions (Day 3). Also, glossing modes were similarly effective for participants' immediate word learning and retention based on the MC test results.

Table 5. Comparing Glossing Modes for Retention in Immediate PR and MC Tests

Test	Day 1	Day 2	Day 3
PR Test	Dual glossing modes (video and audio)	Dual glossing mode (video)	All glossing modes equally effective
MC Test	All glossing modes equally effective	All glossing modes equally effective	All glossing modes equally effective

Different Glossing Modes and Delayed Word Retention

With respect to the participants' performance after two weeks, ANCOVA results showed that Groups B, C, and D were significantly different from the control group on the delayed post-test ($p < .05$) with a small to large effect size for both the PR and the MC tests (.10 to .08 for PR and .14 to .19 for MC). This shows that glossing was partially effective for delayed word learning. To examine which glossing modes were significantly more effective, pairwise comparisons for the PR test (see Table 6) revealed that the three experimental groups gained significantly higher means than the control group ($p < .05$) on all delayed post-tests.

Table 6. Delayed Retention for PR Post-Test

Group	Comparison	Delayed Sub-Test 1		Delayed Sub-Test 2		Delayed Sub-Test 3	
		Difference (M)	p	Difference (M)	p	Difference (M)	p
Group A	Group B	-1.46	.178	-0.32	.999	-1.37	.119
	Group C	-1.65	.127	-3.81*	.000	-1.69*	.030
	Group D	-2.11*	.020	-1.41	.381	-0.37	.988
Group B	Group A	1.46	.178	0.32	.999	1.37	.119
	Group C	-0.19	1.000	-3.48*	.000	-0.31	.993
	Group D	-0.65	.871	-1.09	.544	1.00	.330
Group C	Group A	1.65	.127	3.81*	.000	1.69*	.030
	Group B	0.19	1.000	3.48*	.000	0.31	.993
	Group D	-0.45	.982	2.39*	.008	1.31	.098
Group D	Group A	2.11*	.020	1.41	.381	0.37	.988
	Group B	0.65	.871	1.09	.544	-1.00	.330
	Group C	0.45	.982	-2.39*	.008	-1.31	.098

*The mean difference was significant at $p < .05$

In the Delayed Sub-Test 1 comparison, Group D (L2 definition and video animation glossing) retained more of the target glossed words when compared to Group A ($p < .05$). This result should be interpreted with caution since learners in Group D outperformed those in Group A (control) only. In the second comparison, Group C, with the same glossing mode, outperformed the other three groups on the delayed post-tests ($p < .05$). The third comparison of the Delayed Sub-Test 3 showed that Group C retained the

glossed words to a significant extent only in comparison to the control group ($p < .05$) but not the two other experimental groups. Group C was taught via L2 definition alone.

As for the MC measurement, Table 7 shows that, in the first comparison of the Delayed Sub-Test 1, learners in Group D (L2 definition and video animation glossing) retained words better than participants in Groups A and B ($p < .05$). In the second comparison, learners in Groups C and D retained the words significantly better than participants in Groups A and B ($p < .05$), where L2 definition and video animation and L2 definition alone were used to instruct the words, respectively. In the third comparison, only participants in Group D performed significantly better than Groups A and B ($p < .05$), but not Group C.

Table 7. Delayed Retention for MC Post-Test

Group	Comparison	Delayed Sub-Test 1		Delayed Sub-Test 2		Delayed Sub-Test 3	
		Difference (<i>M</i>)	<i>p</i>	Difference (<i>M</i>)	<i>p</i>	Difference (<i>M</i>)	<i>p</i>
Group A	Group B	-0.10	1.000	-0.80	.783	-0.75	.854
	Group C	-1.50	.454	-2.98*	.000	-2.26*	.015
	Group D	-2.98*	.004	-2.641*	.002	-2.53*	.004
Group B	Group A	0.10	1.000	0.82	.783	0.75	.854
	Group C	-1.39	.309	-2.15*	.005	-1.51	.102
	Group D	-2.88*	.001	-1.82*	.023	-1.78*	.024
Group C	Group A	1.50	.454	2.98*	.000	2.26*	.015
	Group B	1.39	.309	2.15*	.005	1.51	.102
	Group D	-1.48	.337	0.33	.997	-0.26	.999
Group D	Group A	2.98*	.004	2.64*	.002	2.53*	.004
	Group B	2.88*	.001	1.82*	.023	1.78*	.024
	Group C	1.48	.337	-0.33	.997	0.26	.999

*The mean difference was significant at $p < .05$

Table 8 summarizes the findings of glossing modes for both delayed PR and MC tests. As for the PR test, L2 definition and video animation glossing was effective for *two* out of *three* test sessions (Days 1 and 2), whereas L2 definition alone was only influential for *one* out of *three* test sessions (Day 3). In contrast, dual glossing modes were effective for all *three* test sessions of the MC delayed post-test with single glossing mode being positively influential for *two* out of *three* test sessions (Days 2 and 3).

Table 8. Comparing Glossing Modes for PR and MC Delayed Post-Tests

Test	Day 1	Day 2	Day 3
PR Test	Dual glossing mode (video)	Dual glossing mode (video)	Single glossing mode
MC Test	Dual glossing mode (video)	Dual glossing mode (video) and single glossing mode	Dual glossing mode (audio) and single glossing mode

Although the superiority of dual versus single glossing modes is evident for the majority of test sessions in the two vocabulary tests, the findings cannot be generalized with certainty since the single glossing mode was also influential for learners' immediate and delayed word retention in a few test sessions.

Attitudes of Participants Toward the Glossing Modes

To address Research Question 2, a questionnaire was used to examine learners' perceptions of different

glossing modes. Only 83 out of the 132 participants responded to the questionnaire. Table 9 summarizes the findings.

Table 9. *Glossing Modes and Participants' Perceptions*

Questions	L2 Definition and Video Animation Mode		L2 Definition and Audio Mode		Single Mode	
	Percentage	n	Percentage	n	Percentage	n
1–3 It is easy for me to learn new words with this mode	94%	78	76%	63	8%	7
4–6 It was easier for me to remember words in the final test when instructed by this mode	85%	71	64%	53	30%	26
7–9 Given the choice, to learn new words in the future, I would use this mode	91%	76	64%	53	13%	11

Note. Each question was repeated three times for each glossing mode and percentages represent the number of participants' responses who either strongly agreed or agreed to the statement.

Using 9 closed-ended statements, participants were asked to rate their agreement on whether each gloss mode helped with word learning ease, long-term retention, and future practice. As shown in Table 9 (Questions 1–3), participants perceived the dual glossing mode of L2 definition and video animation to be the most beneficial for learning vocabulary (94%), followed by the bimodal glossing of L2 definition and audio (76%), and finally the single glossing mode of L2 definition (8%).

When asked whether instruction using the three glossing modes helped participants retain words (Questions 4–6), 50% of participants strongly agreed and 35% agreed (total 85%) that L2 definition and video animation glossing helped. For text-definition and audio recording, 15% of participants strongly agreed and 49% agreed (total 64%) that it assisted recall. Lastly, 22% agreed and 8% strongly agreed (total 30%) that L2 definitions alone helped them to remember the words easily.

When asked which mode participants would choose to use in the future (Questions 7–9), 65% of participants strongly agreed and 26% agreed that they would use the video animation mode in the future (total 91%). Perceptions toward the use of L2 definition and audio mode were less favorable (64%). The single mode was the least favored mode (13%). In sum, the findings of the questionnaire showed that the participants perceived learning new words via L2 definition and video animation as their most favored glossing mode, because it helped them to learn words better and retain them longer.

The interviews were complimentary in nature and attempted to better capture the participants' voices regarding their preferences for vocabulary learning with the three glossing modes. Overall, the participants preferred the dual modes of L2 definition and video animation glossing and L2 definition and audio glossing over the L2 definition alone. One interviewee stated that watching the video animation clips of the new words along with their simultaneous L2 definitions helped him perform better in the delayed post-test, as he could associate the word meaning with the relevant animation clips: "We keep the images of the words in the mind when we learn them by text and video animation. This strategy helped me a lot to remember the words after two weeks." Another interviewee stated that vocabulary learning via video animation mode was an enjoyable learning experience, which motivated him to learn the target words effectively, and store them in his memory for long-term use. However, another participant favored L2 definition and audio glossing. This aural mode enabled her to learn the correct pronunciation of the new words and gave her a chance to repeat the new words several times for better delayed word retention: "When I know how a word is pronounced, I can learn it better. The audios of the words helped me to remember the definitions fast." In addition, the interviewees reported that they preferred the concurrent presentation of the glossed words

over the single mode, because the dual mode helped them to learn the words in greater depth. Overall, the findings of the interviews indicated that L2 learners perceived the dual mode of L2 definition and video animation glossing to be the most effective for vocabulary learning followed by the L2 definition and audio glossing and the single textual glossing mode.

Discussion

This study examined the efficacy of simultaneous multimedia glossing on L2 learners' immediate and delayed vocabulary learning and retention by drawing upon Mayer's (2014) CTML.

Efficacy of Glossing Modes for Immediate and Delayed Word Learning and Retention

Based on results from the PR test, Table 5 and Table 8 revealed that dual glossing modes were effective for two out of three test sessions in immediate post-tests. All three glossing modes were equally effective for one out of three test sessions. Additionally, all glossing modes were similarly effective for participants' immediate word learning and retention based on the MC test results. Putting together the findings of both immediate and delayed tests, it seems like combining definitions of target words with associated verbal (L2 definition and audio) and visual (L2 definition and video animation) representations is more effective than providing the textual definition of the words alone. The facilitative impact of dual glossing modes was observed in both the PR and MC tests where participants performed significantly better, not only on immediate post-tests, but also partially on delayed tests. This showed a positive effect of glossing modes on learners' immediate and delayed word retention. Results of previous studies on multimedia glossing confirm that multiple glosses are more effective than single or no gloss modes for learners' vocabulary learning (Al-Seghayer, 2001). However, the single glossing mode was also effective for one out of six test sessions in PR, and two out of six test sessions in the MC measurement. Thus, we cannot conclude with certainty that dual modes are more effective than single modes. The next section discusses the efficacy of dual versus single glossing modes.

Efficacy of Dual Glossing Modes

In tune with several studies regarding computer-based multimedia glosses (Abraham, 2008; Akbulut, 2007; Al-Seghayer, 2001; Chun & Plass, 1996a, 1996b; Yeh & Wang, 2003), the superiority of dual over single mode is evident in most of the measurements in this study. The findings also align with the argument by Boers, Warren, Grimshaw, and Siyanova-Chanturia (2017) regarding the amount of attention given to a target word due to several exposures. Dual glossing modes provide more exposure to a target word, and that may explain why learners are more attracted to dual glossing modes than single-glossing modes.

Based on the PR and MC tests, the L2 definition and video animation mode appeared to be significantly more effective than L2 definition and audio glossing. The results of the questionnaires and interviews also showed participants' preferences for the video animation mode. This preference could be a result of learners' limited working memories, preventing them from processing the audio input and leaving them to depend on the visual representation of the video animation to understand the meanings of words (Al-Seghayer, 2016). Another explanation involves Mayer's (2014) definition of multimedia. In his definition, words belong to the verbal mode, whereas pictures belong to the visual category. Using this understanding as the guiding framework, the simultaneous combination of two components (printed text and spoken text) in the verbal form might have induced unnecessary distraction and additional information, resulting in a heavier cognitive load. Mayer and Fiorella (2014) refer to this condition as "redundant multimedia presentation" (p. 287) since learners were exposed to the same target words in both spoken and printed forms. Furthermore, multimedia information might not be effective when verbal or visual channels are overloaded with information of the same category (Sakar & Erçetin, 2005). The explanation sounds plausible for the present study where learners had to process the two verbal modes (text and aural) without receiving any visual input from the eyes (visual channel), resulting in the overloading of one channel. However, with the video animation mode, learners' verbal (L2 definition) and visual (video animation clips) channels were actively challenged, allowing them to receive verbal input from the ears and eyes and visual

input from the eyes. Nevertheless, to minimize the effect of high cognitive load in case of video animation mode, participants received verbal and visual information simultaneously. This discussion may lead some to ponder why the video animation mode is not subject to redundancy and an extraneous load of information for learners. Studies have shown that video animation clips are strong and contextually-based learning instruments that foster vocabulary acquisition and leave learners with a durable learning experience (Al-Seghayer, 2016). Video stimulates learners' curiosity to follow the clips and builds form–meaning relations between a word and its meaning. This form–meaning link is the first condition for meaningful learning to occur (Mayer, 2014), enhancing learners' vocabulary learning and retention.

Efficacy of Single Glossing Modes

The efficacy of the L2 definition alone in a few test sessions of the PR and MC measurements can be accounted for by the following: First, participants might have considered the provision of simple textual glosses as essential for the comprehension of target words (Acha, 2009; Boers, Warren, He, & Deconinck, 2017; Chen, 2016). Second, some of the visual annotations (L2 definition and video animation) might not have been clear enough to understand. For example, an animation clip for the target word *feather*, showing a feather moving in a circle, might represent the feather itself to one learner but the act of spinning to another. Although an important factor, analyzing how video glosses could impede participants' understandings of a word was beyond the scope of this study. Third, cognitive load theory (Chandler & Sweller, 1991) suggests that adding visual representations (e.g., images and videos) to the verbal mode (e.g., textual definitions) may generate a more extraneous load than presenting a verbal annotation alone. Subsequently, the limited working memory of the learners as well as the load of the information may have resulted in their occasional low performance in the sessions that involved dual glossing modes. Fourth, participants' different preferences for verbal or visual representations might have contributed to the efficacy of single versus dual glossing modes in this study. Rassaei (2017) found that learners' preferences to different learning styles affected their vocabulary learning. Attending to the dichotomy of verbalizer–visualizer learning styles would also clarify the question of “for whom is multimedia instruction effective?” (Plass, Chun, Mayer, & Leutner, 1998, p. 25). Fifth, the type of target words to be glossed and the text-type might have affected participants' selection of one glossing mode over others. A meta-analysis (Abraham, 2008) on computer-mediated glosses showed that instruction level, text type, and measurement type are three factors that influence a learner's choice of one gloss type over another. Finally, the type of video animation clips selected to instruct the glossed words might have contributed to the efficacy of a single glossing mode over a dual mode in some test sessions despite piloting all clips twice before the study. Apparently, some clips were not as transparent as they should have been for some target words in one text.

Questionnaire and Semi-Structured Interviews

The findings of this study are consistent with other studies that employed questionnaires and interviews to seek learners' perceptions toward the effectiveness of annotation types for word learning. Al-Seghayer (2001) reported that 86.6% of participants rated video clips as helpful, and in the present study, 94% of the participants rated the video clip as the most helpful mode of vocabulary glossing. Additionally, from the questionnaires and interviews, participants did not perceive the provision of the L2 definition alone as a widely effective gloss for their vocabulary learning and retention. The findings of Al-Seghayer (2001) showed that only 10% of the learners agreed that learning new words would be facilitated if accompanied by the L2 definition alone. Yeh and Wang (2003) reported that participants learned better with pictures than with word definitions alone. Likewise, less than 10% of the learners in this study preferred learning words accompanied by a L2 definition alone. Participants' comments in interviews implied that they perceived textual definition and video animation mode as an interesting and motivating practice for the comprehension of texts in the short term and for delayed retention. However, the inconsistencies between the findings of the PR and MC tests and the questionnaire and interviews still warrant further research.

Limitations and Conclusion

This study faced some limitations. First, the participants had no control over clicking the glossed words in each reading passage. Second, there was inconsistency between methods of instruction and assessment. The participants were exposed to non-traditional hypertexts where the target glossed words were hyperlinked with different multimedia glossing modes, but they were evaluated with a paper-and-pencil method of assessment. Third, because of time restrictions, some videos were cut and edited, resulting in lower comprehensibility of some video clips. Fourth, the delayed post-test was administered only two weeks after the instruction because of study time restraints. Finally, only concrete nouns were selected to gloss, because the video animation clips were able to clearly and easily demonstrate the definitions of these designated words.

Drawing on CTML, this study investigated the effectiveness of simultaneous multimedia glossing on L2 learners' vocabulary learning and retention. The results of the two vocabulary measurements revealed that, in general, dual glossing modes were effective for learners' word learning and retention. However, it could not be certainly stated that dual modes were more effective than the single modes, because the L2 definition alone was influential for learners' vocabulary learning and recollection in a few test sessions. The findings of the questionnaire implied that even though many of the participants favored textual definition and visual annotations as their preferred mode, there were some learners who preferred verbal modes of instruction (i.e., L2 definition alone or L2 definition and audio glossing). Thus, when instructing words via multimedia glossing modes, the learning preferences and styles of all learners in a class are important factors to consider. The interviews also confirmed the preference for the dual mode of L2 definition and video animation glossing. Video animation helped learners retain words for later use and was regarded as an interesting, motivating, and important practice for the comprehension of texts.

References

- Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning*, 21(3), 199–226.
- AbuSeileek, A. (2011). Hypermedia annotation presentation: The effect of location and type on the EFL learners' achievement in reading comprehension and vocabulary acquisition. *Computers & Education*, 57(1), 1281–1291.
- Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology*, 40(1), 23–31.
- Akbulut, Y. (2007). Effects of multimedia annotations on incidental vocabulary learning and reading comprehension of advanced learners of English as a foreign language. *Instructional Science*, 35(6), 499–517.
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202–232. <https://doi.org/10.125/25117>
- Al-Seghayer, K. (2016). Celebrating the story of my first contribution to CALL. *Language Learning & Technology*, 20(2), 179–184. <https://doi.org/10.125/44472>
- Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning*, 17(2), 237–259.
- Boers, F., Warren, P., Grimshaw, G., & Siyanova-Chanturia, A. (2017). On the benefits of multimodal annotations for vocabulary uptake from reading. *Computer Assisted Language Learning*, 30(7), 709–725. <http://doi.org/10.1080/09588221.2017.1356335>
- Boers, F., Warren, P., He, L., & Deconinck, J. (2017). Does adding pictures to glosses enhance vocabulary uptake from reading? *System*, 66, 113–129.

- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition & Instruction*, 8(4), 293–332.
- Chen, I. (2016). Hypertext glosses for foreign language reading comprehension and vocabulary acquisition: Effect of assessment methods. *Computer Assisted Language Learning*, 29(2), 413–426.
- Chun, D. M., & Plass, J. L. (1996a). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal*, 80(2), 183–198.
- Chun, D. M., & Plass, J. L. (1996b). Facilitating reading comprehension with multimedia. *System*, 24(4), 503–519.
- Jacobs, G. M., DuFon, P., & Hong, F. C. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading*, 17(1), 19–28.
- Jung, J. (2016). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research*, 20(1), 92–112.
- Khezlrou, S., & Ellis, R. (2017). Effects of computer-assisted glosses on EFL learners' vocabulary acquisition and reading comprehension in three learning conditions. *System*, 65, 104–116.
- Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Educational Technology and Society*, 11(3), 114–126.
- Kost, C. R., Foss, P., & Lenzini, J. J. (1999). Textual and pictorial glosses: Effectiveness on incidental vocabulary growth when reading in a foreign language. *Foreign Language Annals*, 32(1), 89–113.
- Lin, C., & Tseng, Y. (2012). Videos and animations for vocabulary learning: A study on difficult words. *The Turkish Online Journal of Educational Technology*, 11(4), 346–355.
- Mayer, R. E. (Ed.). (2014). *The Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press.
- Mayer, R. E., & Fiorella, L. (2014). Principles of reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 279–315). New York, NY: Cambridge University Press.
- Mohsen, M. A., & Balakumar, M. (2011). A review of multimedia glosses and their effects on L2 vocabulary acquisition in CALL literature. *ReCALL*, 23(2), 135–159.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge, UK: Cambridge University Press.
- Perez, M. M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning & Technology*, 18(1), 118–141. <https://doi.org/10.125/44357>
- Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Educational Psychology*, 90(1), 25–36.
- Rassaei, E. (2017). Computer-mediated textual and audio glosses, perceptual style, and L2 vocabulary learning. *Language Teaching Research*, 22(6), 657–675. <https://doi.org/10.1177/1362168817690183>
- Rusanganwa, J. A. (2015). Developing a multimedia instrument for technical vocabulary learning: A case of EFL undergraduate physics education. *Computer Assisted Language Learning*, 28(2), 97–111.

- Sakar, A., & Erçetin, G. (2005). Effectiveness of hypermedia annotation for foreign language reading. *Journal of Computer Assisted Learning*, 21, 28–38.
- Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the vocabulary levels test. *Language Testing*, 18(1), 55–88.
- Tabachnick, G. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Vahedi, V. S., Ghonsooly, B., & Pishghadam, R. (2016). Vocabulary glossing: A meta-analysis of the relative effectiveness of different gloss types on L2 vocabulary acquisition. *Teaching English with Technology*, 16(1), 3–25.
- Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, 13(2), 48–67. <https://doi.org/10.125/44180>
- Yeh, Y., & Wang, C. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *CALICO Journal*, 21(1), 131–144.
- Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO Journal*, 20(1), 33–58.
- Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, 24(1), 39–58.
- Webb, S. (2010). Using glossaries to increase the lexical coverage of television programs. *Reading in a Foreign Language*, 22(1), 201–221.

Appendix A. Sample of Different Glossing Modes

Sample Text for Control Group

The Silk Road

In the **ruins** of the ancient Roman city of Pompeii, which was destroyed by a **volcano** in the year 79 C.E., a mirror was found. It had an **ivory** handle in the shape of a female **goddess**. The mirror was from India. In the **tomb** of Li Xian, a Chinese military official who died in 569 C.E., **archeologists** found a water **pitcher** in the shape of a **vase**. The pitcher had a combination of different styles: the shape was from Persia (today's Iran), many details were from central Asia, and the figures on the side were Greek stories from the Trojan War. In the Japanese city of Nara, the 8th century Shosoin Treasures households thousands of exquisite objects of great beauty- furniture, musical instruments, **weapons**, fabric, and military **armor**. These objects come from what is today Vietnam, Western China, Iraq, the Roman Empire, and Egypt. Clearly, long before the globalization of our modern world, trade was going on between very distant lands, and the objects tell a story about a place and time.

Sample Text for L2 Definition Alone Group

In
by a [volcano](#)
the shape
Xian, a Chi
water [pitch](#)
styles: the
Asia, and t
Japanese
of exquisite
fabric, and
Western C
globalization of our modern world, trade was going on between very distant lands,
and the objects tell a story about a place and time.

Volcano
A mountain with a large opening which sends out burned materials.

destroyed
andle in
of Li
nd a
different
n central
r. In the
housands
[pons](#),
am,
e the

Sample Text for L2 Definition and Audio Glossing Group

bought a new one. I tried freezing out my apartment by leaving the door open during the winter chill since I heard the bugs can't survive in temps less than 25 degrees. I
usual apart
reacted to
found the s

Arachnids
A class of insects that include spiders.

Thro
knowing ex
get a new
such as [fle](#)
A bite, whic
made with
cause disc
chemicals
victim's ow
almost all t

ny of the
tem has
e has

ght
, they will
[arachnids](#)
common.
which is
sect bites
ause
bite
t for
easily than

any others. If a dog, a [rodent](#), a horse, or even a snake bites you, you know that it has happened and which creature did it. If an insect or spider bites you, you

Sample Text for L2 Definition and Video Animation Glossing Group

honeybees, and without which they would do badly, is estimated to be about 40,000. This worldwide sea of flowers is pollinated by just nine species and in Europe and
The fact th
for [compe](#)

Ladder
A tool for climbing up and down a wall or a tree.

Tr
negative in
Sichuan in
and [ladder](#)
brush each
transferred
honeybees
Sichuan. P
The US ha
India and C
France, los
government has said bees could completely disappear in less than ten years' time. In recent years, environmentalists have focused on greenhouse gases and

g plants.
little room
tremely
thern
[dusters](#)
usters to
e by
s of
the world.
Brazil,
Europe. In

Appendix B. Part of Reading Passage 1 (*The Silk Road*)

In the ruins of the ancient Roman city of Pompeii, which was destroyed by a volcano in the year 79 C.E., a mirror was found. It had an ivory handle in the shape of a female goddess. The mirror was from India. In the tomb of Li Xian, a Chinese military official who died in 569 C.E., archeologists found a water pitcher in the shape of a vase. The pitcher had a combination of different styles: the shape was from Persia (today's Iran), many details were from central Asia, and the figures on the side were Greek stories from the Trojan War. In the Japanese city of Nara, the 8th century Shosoin Treasures households thousands of exquisite objects of great beauty- furniture, musical instruments, weapons, fabric, and military armor. These objects come from what is today Vietnam, Western China, Iraq, the Roman Empire, and Egypt. Clearly, long before the globalization of our modern world, trade was going on between very distant lands, and the objects tell a story about a place and time.

Appendix C. Part of Reading Passage 2 (*Problems in the Natural World*)

Honeybees are the most important pollinators in most regions of the world where flowering plants exist. However, they are by no means the only insects that play this role. Flies, butterflies, beetles, wasp, bumblebees, and even ants can also pollinate plants. Very few flowers are dependent on a single insect species, although no other pollinators are as effective as are honeybees. In all, 80% of flowering plants worldwide are pollinated by insects and of these about 85% by honeybees. As many as 90% of fruit-trees and flowers are dependent on honeybees. The list of flowering plant pollinated by honeybees includes 170,000 species. The number of flowering plant species that are dependent on honeybees, and without which they would do badly, is estimated to be about 40,000. This worldwide sea of flowers is pollinated by just nine species and in Europe and Africa by only one, which is indispensable for most flowering plants. The fact that honeybees are so successful at pollinating means there is little room for competitors wanting to do the same job. The absence of honeybees from an ecosystem can have an extremely negative impact on human beings. A clear example can be found in southern Sichuan in China. Every year in April, thousands of people take feather, dusters and ladders into the pear orchards and climb the trees. They use the dusters to brush each individual tree in order to collect pollen that will be dried and transferred to other trees. It is a slow and boring job that is normally done by honeybees. More than 20 years ago, pesticides killed all the honeybees of Sichuan. Problems with honeybee populations are occurring all around the world. The US has lost at least 35% of its honeybees in recent years. Canada, Brazil, India and China have also lost huge numbers of bees, as has Western Europe.

Appendix D. Part of Reading Passage 3 (*Bites and Stings*)

The young woman had been looking forward to her nice new apartment in Manhattan. Circumstances turned out to be less comfortable than she expected, as this posting to an online forum about insect bites shows.

I just moved into a newly renovated apartment and got 10 huge, itchy bug bites on my arms, legs, and hip. I thought it was my mattress, so I got rid of it and bought a new one. I tried freezing out my apartment by leaving the door open during the winter chill since I heard the bugs can't survive in temps less than 25 degrees. I went to a dermatologist who said the bite pattern isn't like any of the usual apartment pests, and he didn't know what it was. My immune system has reacted to the bites, and I have prickly itching all over my body. If anyone has found the solution, please email me. Thank you!

Throughout North America, countless people crawl into bed at night knowing exactly how the writer feels. Instead of a peaceful night's sleep, they will get a new round of bites by some mysterious pest. Bites by insects or arachnids such as fleas, ticks, horseflies, mosquitos, or bedbugs are extremely common. A bite, which involves a creature's mouth parts, is different from sting, which is made with a sharp structure appended to a creature's rear end. Most insect bites cause discomfort, if any at all. The bite might cause a little swelling because chemicals in the bug's saliva irritate the skin. When they do happen, the bite victim's

own behavior might be to blame.

Appendix E. Sample of Wordlist Definitions

Archeologist: a person who studies buildings, animals, and objects of the past

Arachnid: a class of insects that include spiders

Armor: metal-covering worn by soldiers to protect the body

Ladder: a tool for climbing up and down a wall or a tree

Volcano: a mountain with a large opening which sends out burned materials

Weapon: an object used for fighting such as a knife, gun, or bomb

Appendix F. Sample of Pre- and Post-PR Vocabulary Test

1. A small insect with a hard covering on its back -----
2. A soft part that covers a bird's body -----
3. A large fly that bites horses and cows -----
4. A doctor who studies and treats skin diseases -----

Appendix G. Sample of Pre- and Post-MC Productive Recognition Vocabulary Test

1. The connection at the top of the leg
 - a. Rib b. Hip c. Knee d. Ankle
2. A woman who is loved, especially for her beauty
 - a. Fairy b. Idol c. Goddess d. Angel
3. A person who buys and sells products
 - a. Shepherd b. Merchant c. Clerk d. Shopkeeper
4. A material used for killing insects
 - a. Trap b. Medicine c. Pesticide d. Fertilizer

Appendix H. Interview Guide

1. How old are you and how long have you been studying English?
2. What strategies do you use to learn vocabulary?
3. What strategies do you use to remember vocabulary?
4. What is your opinion about vocabulary learning through definition-alone?
5. What is your opinion about vocabulary learning through definition and audio pronunciation?
6. What is your opinion about vocabulary learning through definition and video animation?
7. Which vocabulary learning technique (definition alone, definition and audio pronunciation OR definition and video animation) did you prefer/like, and why?
8. Did the definition alone mode help you remember the words in the final test? If no, why not? If yes, in what ways?
9. Did the definition and audio pronunciation mode help you remember the words in the final tests? If no, why not? If yes, in what ways?
10. Did the definition and video animation mode help you remember the words in the final tests? If no,

why not? If yes, in what ways?

11. How have the vocabulary learning modes practiced here changed the way you used to learn new words?
12. What did you like most about the vocabulary learning practice here?
13. What did you like least about the vocabulary learning practice here?
14. What other thoughts do you have to share with me about vocabulary learning by providing text, audio and video definitions?

About the Authors

Nasrin Ramezanali holds a PhD in Applied Linguistics from the University of Western Ontario. Her research interests include L2 and FL acquisition, ESL and EFL, teacher education, educational technology, vocabulary acquisition and learning, and word knowledge measurement.

E-mail: nramezan@uwo.ca

Farahnaz Faez is an Associate Professor in Applied Linguistics in the Faculty of Education at the University of Western Ontario, Canada. Her research interests include L2 teaching and learning, language teacher education, and internationally educated and non-native English speaking teachers.

E-mail: ffaez@uwo.ca